MINDY: AN ALGORITHM FOR THE GENOME-WIDE DISCOVERY OF MODULATORS OF TRANSCRIPTIONAL INTERACTIONS

Wang, Kai [1,2]; Nemenman, Ilya[3]; Banerjee, Nilanjana[2]; Margolin, Adam A.[1,2]; Califano, Andrea[1,2,4]
[1]Department of Biomedical Informatics, Columbia University, New York, NY; [2]Joint Centers for Systems Biology, Columbia University, New York, NY; [3]Los Alamos National Laboratory, Los Alamos, NM; [4]Institute of Cancer Genetics, Columbia University, New York, NY
Keywords: Systems Biology, Regulatory Networks

Transcriptional interactions in the cell are modulated by a variety of mechanisms that prevent their representation as a list of pairwise interactions between transcription factors and their targets. These include, among others, transcription factor activation by phosphorylation and acetylation, formation of active complexes with one or more co-factors, and mRNA/protein degradation and stabilization processes.

This paper presents a first step towards the genome-wide computational inference of genes that modulate the ability of a transcription factor to activate or repress its targets at the post-transcriptional level. Specifically, we developed MINDY, a new information theoretic method to identify multivariate statistical dependencies between a transcription factor and one or more of its targets, conditional on the presence (or absence) of a candidate modulator gene.

MINDY was first validated on a synthetic network model, and then tested in the context of a mammalian cellular network. By analyzing microarray expression profiles from 254 normal and tumor related human B lymphocyte samples, we identified 296 post-transcriptional modulators of the MYC proto-oncogene, responsible overall for modulating 612 regulatory interactions between MYC and its targets. The set is significantly enriched in molecules with function consistent with their activities as modulators of cellular interactions, recapitulates established MYC regulation pathways and known MYC-interacting proteins, and provides a new repertoire of post-transcriptional MYC modulators. Overall, this constitutes the first genome-wide modulation map of a transcription factor.

MINDY has broad applicability and can be used to discover modulators of any other transcription factor, provided that adequate expression profile data are available.

This work was supported by the NCI (1R01CA109755-01A1), the NIAID (1R01AI066116-01), and the National Centers for Biomedical Computing NIH Roadmap initiative (1U54CA121852-01A1). IN was supported in part by DOE/NNSA-LANL LDRD-DR grant No. 20050123DR. AAM is supported by the NLM Medical Informatics Research Training Program (5 T15 LM007079-13).

E-mail: califano@c2b2.columbia.edu